DOCUMENT RESUME

ED 426 703 IR 057 268

AUTHOR	Tescione, Susan M.
TITLE	Invisible Authorship: Women's Names, Databases, and Technology.
PUB DATE	1998-12-29
NOTE	17p.; Revised version of a paper presented at the Annual Conference of Research on Women and Education at Michigan State University (24th, East Lansing, MI, October 22-24, 1998).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Access to Information; *Authors; *Databases; Educational Researchers; *Females; Higher Education; Indexes; *Indexing; Information Dissemination; Information Retrieval; Information Seeking; Information Sources; Internet; *Online Searching; Policy; *Policy Formation; Publications
IDENTIFIERS	ERIC; *Personal Names

ABSTRACT

Bibliographic databases act as search tools to locate relevant literature and information, but they also disseminate information about the works indexed in the records. Articles and authors that cannot be found cannot be cited, and the ability to disseminate data for that particular work is diminished. This study found no significant differences in publication patterns in a small sample of men and women educational researchers. However, results indicated that 50% of the published work of researchers in this sample was not retrievable by author name. The Educational Resources Information Center's (ERIC's) "First Author and Others" policy (which changed in 1998 to include all personal authors appearing on a document) and the Institute for Scientific Information's (ISI's) use of APA (Publication Manual of the American Psychological Association) format in indexing author names limits the ability of users to access records in author name searches. Technology has advanced; these policies are no longer necessary. Core literature indexing policies can be detrimental to women researchers who are more prone to hyphenation, name changes, and name variants than are men. Women are also more likely to be subsequent authors than are men, and are therefore more likely to be limited by database policies. Database policies should be changed to credit all authors for their work. Tips are provided for limiting the damage of the most common pitfalls. (AEF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

IR

Running head: INVISIBLE AUTHORSHIP

ED 426 703

Invisible Authorship:

Women's Names, Databases, and Technology

Susan M. Tescione

University of Washington, Seattle and

Washington Research Institute, Seattle

Presented at the 24th Annual Conference of

Research on Women and Education at Michigan State University

East Lansing, Michigan, October 22-24, 1998

Revised December 29, 1998

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S.M. Tescione

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

157268

Abstract

Bibliographic databases act as search tools to locate relevant literature, but they also disseminate information about the works indexed in the records. This study found no significant differences in publication patterns in a small sample of men and women educational researchers. However, results indicated that 50 percent of the published work of researchers in this sample was not retrievable by author name. ERIC's "First Author and Others" policy* and the Institute for Scientific Information's (ISI) use of APA format in indexing author names limits the ability of users to access records in author name searches. Technology has advanced; these policies are no longer necessary. Core literature indexing limits citation counts, which are used for promotion and tenure review in some academic disciplines. Database indexing policies can be detrimental to women researchers who are more prone to hyphenation, name changes, and name variants than are men. Women are also more likely to be subsequent authors than are men, and are therefore more likely to be limited by database policies. Tips are provided for limiting the damage of the most common pitfalls. Database policies should be changed to credit all authors for their work.

*ERIC's cataloging rules pertaining to Personal Authors changed in 1998, at the same time that ERIC changed and updated its computer software system. Beginning with the January 1998 issue of its abstract journal *Resources in Education* (RIE), ERIC now catalogs *all* Personal Authors appearing on a document. (The old rule was to catalog no more than two authors. If a document cited more than two authors, under the old rule ERIC cataloged only the first name, followed by "And Others").

The bibliographic database acts as a gateway to information about the location of published research. Although the retrieval system locates information for the user, often at the local level, it also disseminates information about the published scholarly work of authors on a global basis. Information concerning the title, author, or contents of the published work that is inaccurate, mistyped, misspelled, or poorly indexed hampers the ability of users to locate that publication (Hamp-Lyons, 1997; Harmon, 1989; Jordan, 1997). Articles and authors that cannot be found cannot be cited, and the ability to disseminate data for that particular work is diminished.

Women often change their names due to marriage custom, and hyphenation merely compounds the difficulty of locating records of their work in library databases, depending upon the programming practices of the databases (Jordan, 1997). Difficulties with name changes, common names, and hyphenation (Courtois & Matthews, 1993) are particularly apparent in the citation indexes.

One common anomaly concerns the frequency with which women's name variants occur in databases. Misspellings, name changes, and nicknames create "splintering" of records within the database, preventing efficient record retrieval (Hamp-Lyons, 1997; Jordan, 1997).

To index articles within journals, large bibliographic databases containing millions of records are compiled by corporations, professional associations, and government agencies. Policies concerning indexing within a database help maintain uniform standards and exercise bibliographic control. Some of the policies in force today are the result of technological limitations that no longer exist and customs that no longer work in the best interests of either the authors of the publications or the people who seek their research.

The ERIC indexing policy of "First Author and Others" in the author field* (ERIC, 1992) and the Institute for Scientific Information's (ISI) policy of indexing cited references by "First Author Only, Last Name and Initials" (DIALOG, 1997; Garfield, 1983) inhibit successful literature searches. ISI indexing is particularly vulnerable to common names, where there are thousands of entries for hundreds of researchers indexed "Smith A," for example. Another ISI policy that impedes dissemination of published work concerns the practice of core literature indexing, which limits how journals are selected for inclusion in the database.

The original purpose of citation indexes was to present a method by which the user could trace both the references and subsequent cites of the article. The user could construct a bibliography of highly relevant material from the most-cited authors on a given topic (Garfield, 1983; Garfield, 1994). In this way, users theoretically had access to records of the top research in any given discipline. As with other bibliographic databases, a search provided the user with a frequency count—in this case, the number of citations that the article had garnered to date from other articles indexed in ISI databases—hence, the term "citation count." By the mid-1970's, academic departments had begun to use citation counts as a measure of quality and quantity of publication (Emmons, 1982). Within ten years, the use of citation counts for the purpose of hiring, tenure, and promotion had become standard procedure in many departments (Garfield, 1983). The citation indexes were frequently used without an understanding of what a citation count was and how it was derived.

In order to be listed in the citation indexes, an author must have published in a journal that was indexed in ISI's "core literature"—a database of approximately 2,000 journals in ISI's top tier. ISI has contractual arrangements with over 8,000 journals, but not all of these journals form the core or top tier (Testa, 1998). The remaining journals form supporting tiers. When a

user makes a query for cites from one of the articles indexed in the core, the records of the cites can come from any of the 8,000 journals indexed. However, if the user requests a citation count for an article from a journal that is not in the core (it could be in one of the supporting tiers, or not in the ISI databases at all), the user receives a return of zero. Records of citations can be obtained only if both the retrieved article and the subsequent cites to that article were made in journals indexed by ISI (Garfield, 1987).

Many faculty misunderstand these concepts. They believe that when a user requests a citation count, the search and retrieval system will return all of the possible citations for the article. Return rates and citation counts of zero are frequent. If the user makes a query in a discipline that is well represented by ISI, the resultant citation count is much more accurate than if the user makes a query in a discipline that is poorly represented in the citation indexes (Garfield, 1987).

Perhaps the most serious result of these policies is the problem of "invisible authorship," a phenomenon that occurs when the authors are not credited for the articles they write or cannot be located within the databases. Since the purpose of the database is to systematically disseminate information about publications, this pilot research began to investigate how database policies and practices can produce the opposite effect. In addition, researchers, particularly women, unknowingly contribute to their own invisible authorship. By analyzing specific cases, techniques for preventing this problem can be examined.

Method

A series of pilot studies were conducted during a one-year period. Names of educational researchers at a large Research I institution were entered into a search and retrieval system to determine methodologies for a larger study. Initial results indicated that faculty at the rank of

assistant professor did not accrue sufficient publication histories to be included in the study. Results for associate professors were inconclusive. Only faculty at the rank of full professor consistently published in sufficient quantity for an extended period. DIALOG databases were selected for the searches.

Participants

In the first portion of the study, eight full professors representing four academic areas in a college of education at a large university were selected for inclusion. Each man was paired to a woman from the same area and where possible, subspecialties, publication history, race, and administrative experience were matched as closely as possible.

This portion of the study represents the lower scoring half of the sample. The four professors represent the two academic areas of Special Education and Leadership and Policy Studies. It must be noted that "lower scoring" for this sample far exceeds national averages for publication and citation counts. The number of publications listed on vitas for this portion of the sample ranged from 35 to 65; of the 181 "searchable" articles, a total of 178 records from multiple databases were retrieved and analyzed for the four participants from the larger body of 1,334 records for the full sample of eight participants. Item-by-item citation counts for the four professors ranged from 33 to 145, when manually added to ensure accuracy. Not all of the publications listed were "indexable"—items that would normally be found in a database. Therefore some publications such as filmstrips and institutional reports were eliminated from the publication lists before the search process began. The number of journal articles actually included in the subset per participant ranged from 30 to 65, averaging 45.25 publications per participant. Three of the four participants began publishing in the 1970's. One participant began publication in 1965.

Instrumentation

Research was conducted at the University of Washington Graduate School of Library and Information Science Computer Laboratory on PC computers with direct connections to DIALOG (DIALOG, 1997) mainframes in California. DIALOG maintains more than 450 databases that are compiled by government agencies, professional associations, and corporate sources. Searches were conducted in the syntactical language developed by DIALOG according to the specifications outlined in the blue sheet technical data.

Seven databases were divided into three groups by software function. Group 1 contained PsycInfo [File 11], Sociological Abstracts [File 37], and the Education Abstracts [File 437]. Group 2 databases, the citation software, included SocialSciSearch [File 7], SciSearch [File 434], and Arts & Humanities Search [File 439]. Note that Group 2 databases index only journal articles, but book reviews and commentaries are indexed as published journal articles in citation software. Participant names were entered into ERIC [File 1] separately.

DIALOG's ONESEARCH feature allowed the operation of all Group 1 databases simultaneously. ONESEARCH was also employed with Group 2 databases. To prevent back reference errors resulting from the use of the remove duplicates (RD) function in citation software, groups of cited references were checked and added manually on an item-by-item basis. Results were recorded on the vita.

Procedures

Unlike the earlier pilots that depended upon global searches, vitas for the four participants were obtained. Separate searches were conducted for each participant, beginning with ERIC, then Group 1 databases, followed by Group 2 databases. The author's name was

entered and results were tallied under the appropriate title in the publication list. The databases containing records of that specific title were noted.

One participant was first author on a number of government reports. Government reports often contain more than one volume, with each volume receiving a separate record. Sometimes sections such as summaries will also be separately indexed. In one case, this author received four records for a government report—two volumes, a summary, and one section of the report had been separately indexed. In cases of multiple indexing like this example, each indexed record denoted a separate publication. Since these reports were listed only in ERIC and not in the citation indexes, this indexing anomaly did not affect citation counts nor did it appear in other databases. Citation counts were conducted, and although they are not essential, the results are summarized in the next section.

Results

This group of researchers was relatively homogenous. T-tests and one-way ANOVA's showed no significant differences between men and women or between the two academic areas, probably due to the small sample. The surprise is in the numbers of records that could not be located. Records not located by author name were added to the number of ERIC documents listed as "and others" in the author field that were not cross-referenced in other databases and are listed as "Invisible" in the last column of Table 1. Table 2 displays similar data in percentages. Table 3 contains a summary of citation indexing.

Table 1: Participant Summary (Raw Data):

Participant	Searchable Publications	Located	Not Located	ERIC "And Others"	Invisible
Man, Sp. Ed.	65	32	33	4	37
Woman, Sp. Ed.	33	16	16	1	17
Man, EDLPS	53	35	18	0	18
Woman, EDLPS	30	14	16	0	16
Total	181	97	83	5	88

Table 2: Participant Summary (Percentages):

Participant	Located	Not Located	ERIC "And Others"	Invisible
Man, Sp. Ed.	49.0 %	51.0 %	6.0 %	57.0%
Woman, Sp. Ed.	48.5 %	48.5 %	3.0 %	51.5 %
Man, LPS	66.0 %	34.0 %	0	34.0 %
Woman, LPS	46.6 %	53.4 %	0	51.2 %

Average lost research for all four authors: 48.4 %

Table 3: Citation Indexing

Participant	No. Records in Citation Index	Percentage of Records in C.I.	Citation Count	No of Average Cites Per Doc.
Man, Sp. Ed.	19	29.2 %	145	13.1
Woman, Sp. Ed.	9	27.3 %	79	8.8
Man, LPS	10	18.9 %	40	4.0
Woman, LPS	12	40.0 %	33	2.8

The approach to this study varies widely from normal procedures for publication patterns. Most studies (Creamer, 1998; Ward, Gast, & Grant, 1989; White, 1985) use the journal as the unit of analysis. The disadvantage to that approach is that only a limited slice of any discipline can be analyzed. When publication patterns for different specialties within a discipline are studied across lengthy careers, subtleties emerge. Creamer (Creamer, 1995) demonstrated this principle in a study in which she interviewed senior women researchers and analyzed their vitas. The number and variety of publications studied is greater and individual case analysis provides clues as to why so much of a researcher's work can be lost in cyberspace. Data analysis also assists in constructing strategies to prevent invisible authorship.

Case 1: ERIC's "First Author and Others" Policy*

By policy, ERIC indexes sole author, first author, and second author in the author name field on each record (ERIC, 1992). The collaborative writing of three authors or more results in the indexing of the first author only with the phrase "and others" replacing the names of all subsequent authors. The policy directly affects women, who are often subsequent authors (Creamer, 1998; Enos, 1990). How often is the "first author and others" policy implemented? AskERIC had the answer. A personal communication noted 112,347 records located with this phrase (C. Sprague, AskERIC Service, September 29, 1998). Each of those records represents at least one author who is named, and a minimum of two who are not listed. It is impossible to estimate the number of researchers this policy affects.

A second problem with this case is the academic area. Both Policy Studies and Special Education appear to be underrepresented in the databases. A number of journals in Special Education were not indexed to databases normally carried in academic libraries. It is entirely possible to publish in a journal that is not indexed in an academic database, or is indexed in a

small professional association database but not to a national database carried by academic libraries. Other researchers might see the article when it is first published, if they happen to read that issue. However, in the case of poor indexing, it is virtually impossible to locate the article in a library search. Problems in this area were noted by Swanson and Trahan (Swanson & Trahan, 1986), but they did not know enough about indexing to conclude that the lack of citations in the subspecialty of learning disabilities could have been created by under-representation and core literature indexing policies (Schwartz, 1997).

Case 2: CD's, Film, and Videotape

Researchers who published in the 1960's and 1970's will find that much of their work does not appear in databases because journals were not extensively indexed in electronic databases at that time. In addition, many researchers do not index or catalog products from their grant projects. These products include videotapes, filmstrips, and computer disks. All of these items can and should be indexed and catalogued.

Case 3: Government Reports

This particular researcher has multiple database records for government reports. Each of those records has been carefully cross-referenced to others from the same report. The documents are Resource in Education records; they contain ED numbers and can be ordered. Some appear on microfiche. The participant in this case was first author, and the documents can be retrieved under his name. However, the subsequent authors for these reports are not listed in the database records. The government reports are indexed in ERIC only and do not appear in other databases.

Case 4: Conference Papers

Conference papers should be indexed. Depending on the institution, ERIC claims that papers it publishes can be considered publications for tenure review (D. P. Ely, Eric

Clearinghouse on Information and Technology, personal communication, September 23, 1998). Indexing conference papers is a simple matter. Most conferences will even distribute forms to encourage presenters to index their papers in ERIC. This process provides the researcher with an opportunity to place the work in a public forum, where many people can access it and other researchers can locate it while searching for related literature. In this case, the researcher did not index conference papers.

Future Research

Plans for future research include adding the item-by-item data from the vitas of the other four researchers in the original sample and adding vitas from other research institutions. Eventually, a minimum of 24 vitas will be analyzed from a variety of academic specialties within education. These vitas will represent more than 1,500 articles published in a wide range of journals. With a larger sample, examining patterns of publication should yield significant results, including data for gender and academic specialties. It should be noted that the four participants that were not included in this segment of the research average 150 articles each and obtain citation counts in the hundreds. When added to the subsample included in this paper, significant differences could emerge.

Implications for Women

There are a number of steps that women researchers can take to limit the damage incurred by database policies. First, all women should publish under one name throughout their entire careers. Unless a user knows the researcher's personal history, it is impossible to locate an entire body of work by a single author if that researcher changes names or uses variants, such as nicknames, especially if common surnames are involved.

Second, many women researchers during the 1970's believed that hyphenation would allow them to use their husbands names and still keep their published work traceable (Foss & Edson, 1989). In some computer languages, hyphens represent the space between words. When a hyphenated name enters a database, the name on the left side of the hyphen is automatically changed to a middle name and depending upon the system, it could become an initial. Thus Kate Smith-Jones becomes Jones, KS in some databases. In other databases, the hyphen is stripped, resulting in Smithjones, K. Still other databases retain the left side of the hyphen as the middle name, indexed as Jones, Kate Smith or sometimes Jones, Kate S. As many as eight name variations were located during the pilot studies for some women. In contrast, most men researchers have only two variations—their full names with initials and without initials. Using the name expansion feature on most databases overcomes this relatively minor splintering for men.

The third problem concerns journal indexing policies that women can work to overcome. Conducting research and publishing consumes an enormous amount of time and effort. To obtain a return on their investment, women should publish their work in the highest quality journals. Journals that cross-index to several national databases are more desirable than journals that are not refereed and not indexed. There are two quick methods to determine cross-indexing. The ISI databases appear to have a high correlation with cross-indexing. Therefore, women can telnet to ISI's website and conduct a title word search of the Master List (ISI, 1998) to determine which journals are indexed by ISI. The name of the journal can then be input into databases at academic libraries to determine if other databases index that journal. If a journal from ISI's Master List cross-references to several databases, consider that journal as a source for publication. A reference librarian can also facilitate this process with access to DIALOG's Journal Name Finder

[File 414] (DIALOG, 1995). The librarian can input the name of the journal, and the results will yield the names of the databases that index that journal. Some journals list the databases that index their articles in each issue. Check the journal's indexing before submitting an article for publication.

Index conference papers and grant project products such as CD's, videotapes, and other media. These are opportunities for women to put their names in databases, where information about these products can be disseminated.

One useful strategy is to encourage women students to publish and to present papers at conferences while still in graduate school. For women to be taken seriously in the current academic climate, they need to be more aware of what should be accomplished beyond their classroom assignments. All eight of the full professors in the original sample for this study began publishing as graduate students. The faculty with the fewest publications in this subset produce nearly triple the national average for numbers of publications (Creamer, 1998). It is not enough to demonstrate that patterns of discrimination exist in academe, particularly in hiring, promotion, and tenure review (Hornig, 1980). If women wish to enter the ranks of tenured faculty, they must meet the standards of the academic culture. The academy changes slowly, and no one denies that there is a glass ceiling. The emphasis on publication has become more pronounced in recent years—not less so. The trend is unlikely to reverse itself, and women researchers must respond accordingly.

*Note

ERIC indexing policy has recently changed. In 1998, ERIC began to index all authors (T. Brandhorst, ERIC Processing and Reference Facility, personal communication, December 18, 1998, in reference to ERIC internal memo, February 19, 1998).

References

Courtois, Martin P. & Matthews, Judith A. (1993). Tips for searching ISI citation indexes for personnel decisions. *Database, 16*, 60-67.

Creamer, Elizabeth G. (1995). The scholarly productivity of women academics. *Initiatives, 57*, 1-9.

Creamer, Elizabeth G. (1998). Assessing faculty publication productivity, *Ashe-ERIC Higher Education Report* (Vol. 26, pp. 1-117). Washington, DC: George Washington University Graduate School of Education and Human Development.

DIALOG. (1995). *DIALOG pocket guide*. Mountain View, CA: Knight-Ridder.

DIALOG. (1997). *Bluesheets*, [Online]. Knight Ridder. Available: http://library.dialog.com/bluesheets [1997, .

Emmons, Carol-Ann. (1982). A longitudinal study of the careers of a cohort of assistant professors in psychology. *American Psychologist, 37*, 1228-1238.

Enos, Theresa. (1990). Gender and publishing. *Pre-Text, 11*, 311-316.

ERIC. (1992). *Database master files tape documentation* (Systems Manual). Washington, DC: ERIC Processing and Reference Facility.

Foss, Karen A. & Edson, Belle A. (1989). What's in a name? Accounts of married women's name choices. *Western Journal of Speech Communication, 53*(Fall 1989), 356-373.

Garfield, Eugene. (1983). How to use citation analysis for faculty evaluations, and when is it relevant? *Current Comments, Parts I and II*(44-45), 354-372.

Garfield, Eugene. (1987). Citation data is subtle stuff. *Scientist, 1*(6).

Garfield, Eugene. (1994). *The concept of citaton indexing: A unique and innovative tool for navigating the research literature*, [Online]. Institute for Scientific Information. Available: http://wwww.isnet.com/essays/essay1.html.

Hamp-Lyons, Liz. (1997). More thoughts on academic naming practices. *English for Specific Purposes, 16*, 73-74.

Harmon, Robert B. (1989). *Elements of bibliography: A simplified approach.* Metuchen, NJ: Scarecrow Press.

Hornig, Lilli S. (1980). Untenured and tenuous: The status of women faculty. *Annals of the American Academy of Political Science, 448*(March, 1980), 115-125.

ISI. (1998). *Journal Master List Links*, [On line]. Available: http://www.isinet.com/listlinx.html.

Jordan, R. R. (1997). What's in a name? Consistency! *English for Specific Purposes, 16*, 71-72.

Schwartz, Charles A. (1997). The rise and fall of uncitedness. *College & Research Libraries, 58*, 19-29.

Swanson, H. Lee & Trahan, Marcy. (1986). Characteristics of frequently cited articles in learning disabilities. *Journal of Special Education, 20*, 167-182.

Testa, James A. (1998). *The ISI database: The journal selection process*, [Online]. Institute for Scientific Information. Available: http://isinet.com/whatshot/essays/9701.html.

Ward, Kathryn B., Gast, Julie, & Grant, Linda. (1989). Visibility and dissemination of women's and men's sociological scholarship. *Social Problems, 39*, 291-298.

White, Arden. (1985). Women as authors and editors of psychological journals. *American Psychologist, 40*, 527-530.